

Rehaussement de la parole multicanal pour patients atteints de neuropathies auditives

Lieu : [LORIA](#), équipe [MULTISPEECH](#), Nancy

Encadrants : Romain Serizel (Maître de Conférences, Université de Lorraine), Paul Magron (Chargé de Recherche, INRIA).

Cette thèse s'inscrit dans le cadre du projet ANR 'REFINED' impliquant l'équipe Multispeech du Laboratoire lorrain de recherche en informatique et ses applications ([LORIA](#)) à Nancy, le laboratoire de d'intelligence artificielle embarquée au CEA ([List](#)) à Paris et l'[Institut de l'audition](#) à Paris.

Contexte

Dans le monde, environ 466 millions de personnes souffrent actuellement d'une perte auditive. Pour remédier à la perte de sensibilité auditive, depuis près d'un siècle, on conçoit des aides auditives portables. Malgré les progrès récents dans le traitement du signal audio intégré dans les modèles d'aides auditives actuels, les personnes souffrant de neuropathies auditives bénéficient de peu ou pas d'avantages des aides auditives actuelles [1]. Contrairement aux pertes auditives habituelles, les pertes liées à une neuropathie auditive altèrent le traitement de l'information temporelle sans nécessairement affecter la sensibilité auditive. Cela peut avoir un impact particulièrement important dans les scénarios où la parole d'intérêt est présente avec un bruit de fond ou avec plusieurs locuteurs simultanés.

Les systèmes de rehaussement de la parole actuels sont généralement entraînés sur des jeux de données génériques, et ils sont conçus pour optimiser un coût générique, comme l'erreur quadratique moyenne [2] ou le rapport parole/distorsion [3] entre la parole cible (connue) et la sortie du système, qui est estimée à partir du mélange. Le système est ensuite évalué à l'aide d'un critère conçu pour refléter la perception de la parole par des personnes sans perte auditive [4]. Or, le besoin principal des sujets atteints de neuropathies auditives, partagé avec les sujets vieillissants qui éprouvent des difficultés de traitement auditif central, n'est pas de restaurer l'audibilité mais d'améliorer leur intelligibilité de la parole, en particulier dans les environnements bruyants, en compensant la détérioration des signaux acoustiques, qui reposent sur la précision temporelle [5].

Objectifs

Sur la base d'études cliniques réalisées par l'Institut de l'Audition dans le cadre du projet, l'objectif principal de cette thèse est de définir de nouvelles fonctions de coût à optimiser par les algorithmes de traitement de la parole qui soient plus pertinentes pour les sujets atteints de neuropathie auditive que les coûts génériques utilisés dans les algorithmes actuels. Nous porterons une attention particulière à la capacité des algorithmes à aider les volontaires dans des scénarios avec de multiples sources cibles potentielles réparties spatialement dans une pièce. Nous concevrons des filtres d'amélioration de la parole visant à extraire non seulement la parole, mais également des indices supplémentaires tels que l'enveloppe de la parole ou le timbre. Dans une dernière étape, le modèle sera adapté avec une supervision humaine minimale afin de réduire l'étape contraignante

des ajustements itératifs 'manuels' habituels lors de visites répétées avec un clinicien spécialiste pour adapter l'aide auditive aux besoins individuels.

Profil

- Profil solide en traitement du signal audio ou en apprentissage automatique
- Excellentes compétences en programmation
- Excellentes compétences en anglais écrit

Candidature

Déposez votre dossier de candidature sur le site ADUM

(https://www.adum.fr/as/ed/voirproposition.pl?site=adumR&matricule_prop=43498#version) avec les pièces suivantes :

- CV
- Lettre de motivation
- Lettre de recommandation
- Relevé de notes M1-M2
- Mémoire de master, si disponible

Références

- [1] Berlin, C. I. et al. Multi-site diagnosis and management of 260 patients with auditory neuropathy/dys-synchrony (auditory neuropathy spectrum disorder). *Int J Audiol* 49, 30-43 (2010).
- [2] Doclo, S., Spriet, A., Wouters, J. & Moonen, M. Frequency-domain criterion for the speech distortion weighted multichannel Wiener filter for robust noise reduction. *Speech Communication* 49, 636-656 (2007).
- [3] Luo, Y., et al. FaSNet: Low-latency adaptive beamforming for multi-microphone audio processing. 2019 IEEE automatic speech recognition and understanding workshop (2019).
- [4] Vincent, E., Rémi G., and Cédric F. Performance measurement in blind audio source separation. *IEEE transactions on audio, speech, and language processing* 14.4, 1462-1469 (2006).
- [5] https://claritychallenge.github.io/clarity_CC_doc/docs/cpc1/cpc1_intro

Multichannel Speech Enhancement for Patients with Auditory Neuropathy Spectrum Disorders

Location: [LORIA](#), [MULTISPEECH](#) team, Nancy

Supervisors: Romain Serizel (Maître de Conférences, Université de Lorraine), Paul Magron (Chargé de Recherche, INRIA).

This PhD fits within the scope of the ANR project 'REFINED' involving the Multispeech research team in ([LORIA](#)), Nancy, the Laboratory of Embedded Artificial Intelligence in CEA ([List](#)) in Paris, and the [Hearing Institute](#) in Paris.

Context

Worldwide, around 466 million people currently suffer from a hearing loss. To remedy the loss of hearing sensitivity, portable hearing aids have been designed for almost a century. Regardless of the recent advances in audio signal processing integrated in current hearing aids models, people suffering from Auditory Neuropathy Spectrum Disorders enjoy little or no benefit from current hearing aids [1]. Contrary to regular hearing losses, Auditory Neuropathy Spectrum Disorders impair the processing of temporal information without necessarily affecting auditory sensitivity. This can have a particularly dramatic impact in scenarios where the speech of interest is present together with some background noise or with one or several concurrent speaker(s).

Current speech enhancement systems are usually trained on generic corpora, and they are designed to optimize some cost between the target (known) speech and the output of the system, which is estimated from the mixture, such as the mean squared error [2] or the speech-to-distortion ratio [3]. The trained system is then evaluated using a criterion that is designed to reflect the speech perception from people without hearing losses [4]. Yet, the main need of subjects with Auditory Neuropathy Spectrum Disorders, shared with ageing subjects who experience central auditory-processing difficulties, is not to restore audibility but to improve their speech intelligibility, particularly in noisy environments, by compensating for the deterioration of acoustic cues that rely on temporal precision [5].

Objectives

Based on clinical studies performed at the Hearing Institute within the project, the main goal of this PhD is to define new cost functions to be optimized by the speech processing algorithms that are more relevant for subjects with Auditory Neuropathy Spectrum Disorders than generic losses used in current algorithms. We will pay particular attention to the algorithms' ability to help volunteers in scenarios with multiple potential target sources that are spatially distributed in a room. We will derive the speech enhancement filters aiming to extract not only speech, but also additional cues such as speech contour or timbre. In a latter step, the model will be adapted under light human supervision in order to reduce the burden of the usual iterative "handcrafted" adjustments and repeated visits with a specialist clinician to fit the hearing aid to individual needs.

Profile

- Strong background in audio signal processing or machine learning

- Excellent programming skills
- Excellent English writing and speaking skills

Application

Upload your application on ADUM (https://www.adum.fr/as/ed/voirproposition.pl?site=adumR&matricule_prop=43498#version) with the following:

- CV
- Cover letter
- Recommendation letter
- M1-M2 note transcript
- Master thesis, if available

References

- [1] Berlin, C. I. et al. Multi-site diagnosis and management of 260 patients with auditory neuropathy/dys-synchrony (auditory neuropathy spectrum disorder). *Int J Audiol* 49, 30-43 (2010).
- [2] Doclo, S., Spriet, A., Wouters, J. & Moonen, M. Frequency-domain criterion for the speech distortion weighted multichannel Wiener filter for robust noise reduction. *Speech Communication* 49, 636-656 (2007).
- [3] Luo, Y., et al. FaSNet: Low-latency adaptive beamforming for multi-microphone audio processing. 2019 IEEE automatic speech recognition and understanding workshop (2019).
- [4] Vincent, E., Rémi G., and Cédric F. Performance measurement in blind audio source separation. *IEEE transactions on audio, speech, and language processing* 14.4, 1462-1469 (2006).
- [5] https://claritychallenge.github.io/clarity_CC_doc/docs/cpc1/cpc1_intro